

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 02-064832

(43)Date of publication of application : 05.03.1990

(51)Int.Cl.

G06F 12/08

G06F 3/06

G06F 12/12

(21)Application number : 63-215204

(71)Applicant : NEC CORP

(22)Date of filing : 31.08.1988

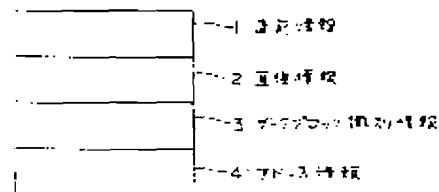
(72)Inventor : NANBU SHUNEI

(54) DISK CACHE MEMORY CONTROL DEVICE

(57)Abstract:

PURPOSE: To prevent an adverse influence from being given to the performance of a whole computer system by sorting a disk cache memory area in the order of the high frequency of access with respect to a data block stored in the area.

CONSTITUTION: A chain of control information is constituted by using two pieces of information in a position on the main storage device of the computer system in a control information area corresponding to the disk cache memory which is precedingly positioned in the order relation of the access frequency that is stored in a storage area 1, and a position on the main storage device which is succeedingly positioned by the order relation of the access frequency that is stored in a storage area 2. The chain is controlled so that the data blocks stored in the corresponding areas are sorted in a descending order from the high frequency of access to the low frequency. Thus, the adverse influence is prevented from being given to the performance of the whole computer system.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's

decision of rejection]

[Date of extinction of right]

⑩ 日本国特許庁(JP)

⑪ 特許出願公開

⑫ 公開特許公報(A) 平2-64832

⑬ Int. Cl.⁵

G 06 F 12/08
3/06
12/12

識別記号

3 2 0
3 0 2 A
A

庁内整理番号

7010-5B
6711-5B
7010-5B

⑭ 公開 平成2年(1990)3月5日

審査請求 未請求 請求項の数 1 (全5頁)

⑮ 発明の名称 ディスク・キャッシュ・メモリ管理装置

⑯ 特 願 昭63-215204

⑰ 出 願 昭63(1988)8月31日

⑱ 発 明 者 南 部 俊 英 東京都港区芝5丁目33番1号 日本電気株式会社内

⑲ 出 願 人 日本電気株式会社 東京都港区芝5丁目33番1号

⑳ 代 理 人 弁理士 渡 辺 喜 平

明 細 書

1. 発明の名称

ディスク・キャッシュ・メモリ管理装置

2. 特許請求の範囲

個々のディスク・キャッシュ・メモリ領域の管理情報に、当該データ・ブロックのアクセス頻度に応じた順序関係を表す情報を格納するアクセス頻度順序関係情報格納手段と、データ・ブロックのアクセス頻度に応じて管理情報領域に格納されているアクセス頻度順序関係を更新するアクセス頻度順序関係情報更新手段と、ディスク・キャッシュ・メモリ上の全ての領域が使用中である状況でディスク・キャッシュ・メモリ上に格納されていない新たなデータ・ブロックへのアクセスが発生した場合、ディスク・キャッシュ・メモリ領域の中での順序関係において最低位のものを選択し、データ・ブロックを実際に読み込んだ後、順序関

係を最低位として管理情報領域に格納する領域選択手段とを具備することを特徴とするディスク・キャッシュ・メモリ管理装置。

3. 発明の詳細な説明

〔産業上の利用分野〕

本発明は、ランダムアクセス可能な低速補助記憶装置におけるアクセスを高速化せしめるディスク・キャッシュ・メモリ管理装置に関し、特に、計算機システムにおける処理速度をより高速化せしめるディスク・キャッシュ・メモリ管理装置に関する。

〔従来の技術〕

補助記憶装置である磁気ディスク装置などは、一回のデータのアクセスを特定の大きさを持つ単位にブロック化して行ない、かつ、記憶媒体上の任意のデータブロックをランダムにアクセスすることが可能であるが、アクセスの速度は低速である。このため、このような低速な記憶媒体を使用

した計算機システムでは、その補助記憶装置に格納されている任意のデータ・ブロックに対してランダムに発生するアクセスを高速化するため、ディスク・キャッシュ・メモリが利用されている。

このディスク・キャッシュ・メモリは、計算機の持つ主記憶装置のうちの特定の領域に、有限個数の補助記憶装置へのデータ・アクセス単位に対応したデータ・ブロックに格納されているデータを常に保持させておき、補助記憶装置に格納されているデータ・ブロックに対するアクセスが発生した場合、補助記憶装置への実際のアクセスに先立ち、この主記憶装置上の特定領域に保持されているデータ・ブロックの中に必要とするデータ・ブロックが存在しているか否かを検査し、存在している場合は補助記憶装置に対するアクセスを行わずに主記憶装置上の特定領域に保持しているデータ・ブロックに対するアクセスを行なうものである。

メモリ領域に対応する管理情報領域に格納されているキャッシュ・メモリに格納されているデータ・ブロックを補助記憶装置からディスク・キャッシュ・メモリ領域に読み込んだ時刻と、そのデータ・ブロックに対するアクセス回数の二つの情報をもとに、該当するデータ・ブロックのアクセス頻度を算出し、さらに、全てのディスク・キャッシュ・メモリ領域の中から算出されたアクセス頻度が最も低いデータ・ブロックを格納している領域を選び出し、その領域に新たなブロックを格納していた。

[解決すべき問題点]

上述した従来のディスク・キャッシュ・メモリ管理装置は、ディスク・キャッシュ・メモリ上の全ての領域が使用中である状況でディスク・キャッシュ・メモリ上に格納されていない新たなデータ・ブロックへのアクセスが発生した場合、そのデータ・ブロックを格納するためにディスク・キ

ところで、従来のディスク・キャッシュ・メモリ管理装置は、一つ一つのデータ・ブロックに対応するキャッシュメモリ領域に対応した管理情報領域を設け、その中に対応するキャッシュ・メモリに格納されているデータ・ブロックが補助記憶装置から読み込まれた時間を設定している。そして、その後、該当するデータ・ブロックに対するアクセスが発生するたびにその管理情報領域の中に格納されているアクセス回数の更新を行ない、常に管理情報領域の中に該当するデータ・ブロックに対するアクセスの頻度を算出するための基本的情報を保持している。

また、ディスク・キャッシュ・メモリ上の全ての領域が使用中であるときにディスク・キャッシュ・メモリ上に格納されていない新たなデータ・ブロックへのアクセスが発生した場合は、この新たなデータ・ブロックを格納するための領域を選択するにあたり、個々のディスク・キャッシュ・

キャッシュ・メモリ領域を選択するにあたり、全てのディスク・キャッシュ・メモリ領域に対応する管理情報領域を検索し、そこに格納されている個々のデータ・ブロックに対するアクセス回数とそのデータ・ブロックを補助記憶装置からディスク・キャッシュ・メモリ領域へ読み込んだ時刻をもとにそのデータ・ブロックへのアクセス頻度を算出する必要があった。しかし、全てのディスク・キャッシュ・メモリ領域に対応する管理情報領域の探索は、ディスク・キャッシュ・メモリ領域の個数が多くなれば多くなるほどオーバーヘッドが増大するため、計算機システム全体の性能に対して悪影響を与えるという問題点があった。特に、マイクロ・プロセッサなどの比較的処理能力の低い中央演算処理装置を使用した計算機システムに対しては顕著である。

また、上述したように時刻情報を使用していたため、時計を持たないような小規模な計算機シス

テムでは採用できないという問題点もあった。

本発明は、上記問題点にかんがみてなされたもので、ディスク・キャッシュ・メモリ領域の個数が多くなった場合でも、計算機システム全体の性能に対して悪影響を与えることのないディスク・キャッシュ・メモリ管理装置の提供を目的とする。

〔問題点の解決手段〕

上記目的を達成するため、本発明のディスク・キャッシュ・メモリ管理装置は、個々のディスク・キャッシュ・メモリ領域の管理情報に、当該データ・ブロックのアクセス頻度に応じた順序関係を表す情報を格納するアクセス頻度順序関係情報格納手段と、データ・ブロックのアクセス頻度に応じて管理情報領域に格納されているアクセス頻度順序関係を更新するアクセス頻度順序関係情報更新手段と、ディスク・キャッシュ・メモリ上の全ての領域が使用中である状況でディスク・キャッシュ・メモリ上に格納されていない新たなデー

タ・ブロックへのアクセスが発生した場合、ディスク・キャッシュ・メモリ領域の中での順序関係において最低位のものを選択し、データ・ブロックを実際に読み込んだ後、順序関係を最低位として管理情報領域に格納する領域選択手段とを備えた構成としてある。

かかる構成とすることにより、ディスク・キャッシュ・メモリ上の全ての領域が使用中である状況でディスク・キャッシュ・メモリ上に格納されていない新たなデータ・ブロックへのアクセスが発生した場合、一つ一つのデータ・ブロックが格納されている個々のディスク・キャッシュ・メモリ領域相互の間に、格納されているデータ・ブロックに対するアクセスの頻度にもとづいて順序関係をアクセス時刻など、そのデータ・ブロックがディスク・キャッシュ・メモリ上に格納されてから経過している時間を定量的に表す情報を用いずに設け、一回のデータ・ブロックに対するアクセ

スのたびにこの順序関係を更新し、ディスク・キャッシュ・メモリ上の全てのディスク・キャッシュ・メモリ領域をそこに格納されているデータ・ブロックに対するアクセスの頻度の高い順序でソーティングしておき、ディスク・キャッシュ・メモリ上の全ての領域が使用中である状況でディスク・キャッシュ・メモリ上に格納されていない新たなデータ・ブロックへのアクセスが発生した場合にそのデータ・ブロックを格納するためのディスク・キャッシュ・メモリ領域として、ソーティングしてある並びの最後尾のものを選択する。

〔実施例〕

以下、図面にもとづいて本発明の実施例を説明する。

第1図は、本発明の一実施例に係るディスク・キャッシュ・メモリ管理装置におけるディスク・キャッシュ・メモリ領域に対応する管理情報領域の構成図である。

同図において、1はアクセス頻度の順序関係で直前に位置づけられるディスク・キャッシュ・メモリ領域に対応する管理情報領域の計算機システムの主記憶装置上での位置を表す情報（直前情報）の格納領域、2はアクセス頻度の順序関係で直後に位置づけられるディスク・キャッシュ・メモリ領域に対応する管理情報領域の計算機システムの主記憶装置上での位置を表す情報（直後情報）の格納領域、3は対応するディスク・キャッシュ・メモリ領域に格納されているデータ・ブロックを識別する情報（データブロック識別情報）の格納領域、4は対応するディスク・キャッシュ・メモリ領域の計算機システムの主記憶装置上での位置を表す情報（アドレス情報）の格納領域である。

格納領域1に格納されるアクセス頻度の順序関係で直前に位置づけられるディスク・キャッシュ・メモリ領域に対応する管理情報領域の計算機システムの主記憶装置上での位置と、格納領域2に

格納されるアクセス頻度の順序関係で直後に位置づけられるディスク・キャッシュ・メモリ領域に対応する管理情報領域の計算機システムの主記憶装置上での位置の二つの情報を使用して、これらの管理情報の連鎖を構成し、この連鎖をそれぞれに対応するディスク・キャッシュ・メモリ領域に格納されているデータ・ブロックのアクセス頻度に応じて常にソーティングされるように管理を行なう。

次に、第2図は、第1図のディスク・キャッシュ・メモリ管理装置における連鎖によるディスク・キャッシュ・メモリ管理情報を示す図である。

同図において、5～9はディスク・キャッシュ・メモリ領域に対応する管理情報であり、それぞれアクセス頻度の順序関係の高いものから低いものへ連鎖している。

初期状態においては、この管理情報を構成しているディスク・キャッシュ・メモリ領域に対応す

て最高位の管理情報5に位置する場合は、単にそのディスク・キャッシュ・メモリ領域をアクセスするものとする。

一方、必要とするデータ・ブロックを格納しているディスク・キャッシュ・メモリ領域に対応する管理情報領域がディスク・キャッシュ・メモリ管理情報の6～9に位置する場合は、そのディスク・キャッシュ・メモリ領域をアクセスするとともに、その管理情報のディスク・キャッシュ・メモリ管理情報での位置を他の位置のものと交換するべく、第1図における格納領域1および2の内容を更新する。ディスク・キャッシュ・メモリ上に格納されているデータ・ブロックに対するアクセスが発生するたびにこの処理を行なうことにより、ディスク・キャッシュ・メモリ管理情報5～9を構成するディスク・キャッシュ・メモリ領域に対応する管理情報領域の連鎖は、それぞれの管理情報領域に対応するディスク・キャッシュ・メ

る管理情報領域は存在しない。この管理情報は、ディスク・キャッシュ・メモリ上の全ての領域が使用中になった時点で全てのディスク・キャッシュ・メモリ領域に対応する管理情報領域が含まれるようになるものとする。この状態に至るまでの間に発生した新たなデータ・ブロックに対するアクセスの処理については、空きであるディスク・キャッシュ・メモリ領域を使用し、その領域に必要なデータ・ブロックを補助記憶装置から読み込み格納した後、その領域に対応する管理情報領域にデータを設定し、その管理情報を管理情報9で示される位置に追加する。

ディスク・キャッシュ・メモリ上に格納されているデータ・ブロックに対するアクセスが発生した場合は、ディスク・キャッシュ・メモリ領域の空きの有無にかかわらず、必要とするデータ・ブロックを格納しているディスク・キャッシュ・メモリ領域に対応する管理情報領域が管理情報の中

に格納されているデータ・ブロックに対するアクセスの頻度が高いものから低いものへ降順にソーティングされる。

また、ディスク・キャッシュ・メモリ領域が全て使用中である状況で、ディスク・キャッシュ・メモリ上に格納されていない新たなデータ・ブロックに対するアクセスが発生した場合、ディスク・キャッシュ・メモリ管理情報5～9は上述の通りすでにアクセス頻度順でソーティングされているため、新たに必要となったデータ・ブロックを格納するためのディスク・キャッシュ・メモリ領域としては、ディスク・キャッシュ・メモリ管理情報5～9の中で最もアクセス頻度の低いデータ・ブロックを格納している管理情報9の位置にあるディスク・キャッシュ・メモリ領域管理情報に対する領域を容易に選択することが可能である。

このように本実施例では、ディスク・キャッシュ方式におけるディスク・キャッシュ・メモリの

管理方式に、一つ一つのデータ・ブロックが格納されている個々のディスク・キャッシュ・メモリ領域相互の間に、格納されているデータ・ブロックに対するアクセスの頻度にもとづいた順序関係を設け、一回のデータ・ブロックに対するアクセスのたびにこの順序関係を更新し、ディスク・キャッシュ・メモリ上の全てのディスク・キャッシュ・メモリ領域をそこに格納されているデータ・ブロックに対するアクセスの頻度の高い順序でソーティングしておき、ディスク・キャッシュ・メモリ上の全てのディスク・キャッシュ領域が使用中である状況でディスク・キャッシュ・メモリ上に格納されていない新たなデータ・ブロックへのアクセスが発生した場合に、そのデータ・ブロックを格納するためのディスク・キャッシュ・メモリ領域として、ソーティングしてある並びの最後尾のものを選択する方式を採っている。

5～9：アクセス頻度順のディスク・
キャッシュ・メモリ管理情報

代理人 弁理士 渡辺 喜平

〔発明の効果〕

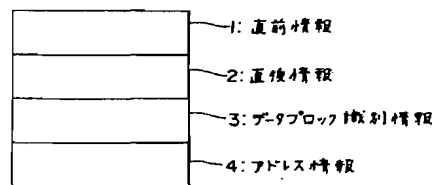
以上説明したように本発明は、従来、複雑非効率的なディスク・キャッシュ・メモリ管理処理そのものを簡素化できる上、比較的处理能力の低い中央演算処理装置を使用した小規模な計算機システムにおいても効率の良いディスク・キャッシュ方式によって補助記憶装置のアクセスの高速化を図ることが可能なディスク・キャッシュ・メモリ管理装置を提供できるという効果がある。

4. 図面の簡単な説明

第1図は本発明の一実施例に係るディスク・キャッシュ・メモリ管理装置におけるディスク・キャッシュ・メモリ領域に対応する管理情報領域の構成図、第2図は第1図のディスク・キャッシュ・メモリ管理装置における連鎖によるディスク・キャッシュ・メモリ管理情報を示す図である。

1～4：格納領域

第 1 図



第 2 図

